THE FIFTH IEEE CONFERENCE ON
MULTIMEDIA BIG DATA (BigMM-19)

Photo Filter Recommendation Through Analyzing Objects, Scenes and Aesthetics

Yi-Ning Chen and **Mei-Chen Yeh**

Dept. of Computer Science and Information Engineering

National Taiwan Normal University

Social Media for Photo Sharing

- Monthly actives 800M+
- Daily posts 40M+
- Half of photos are filtered.

Photo Filter

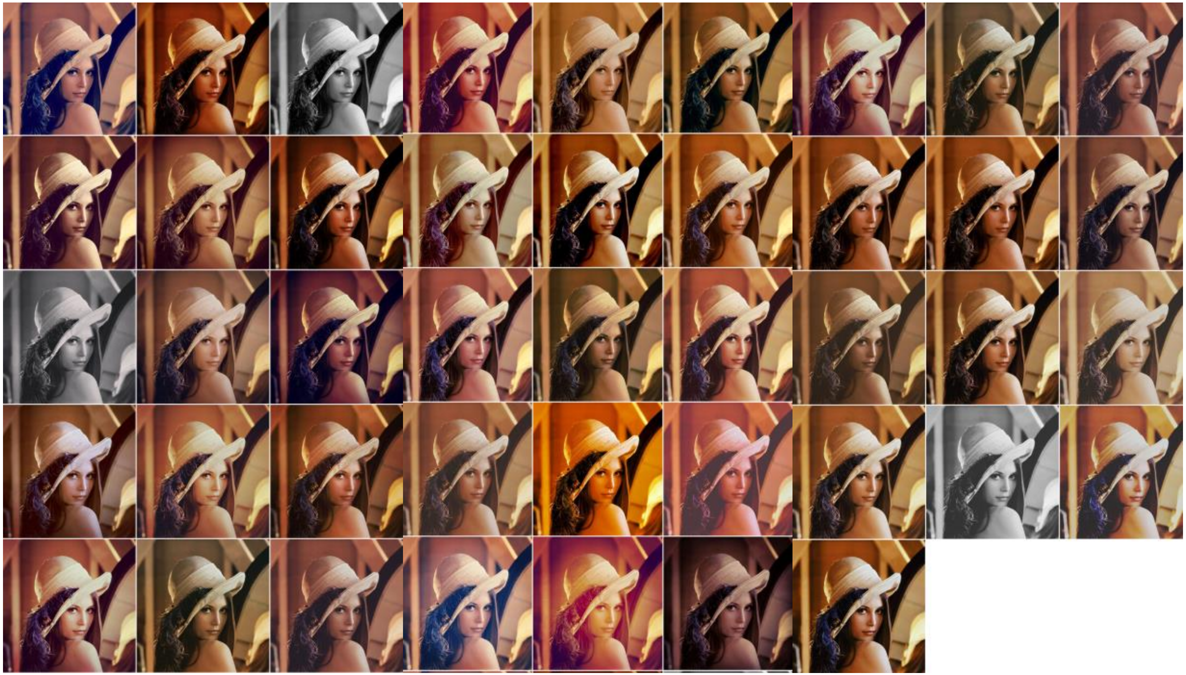
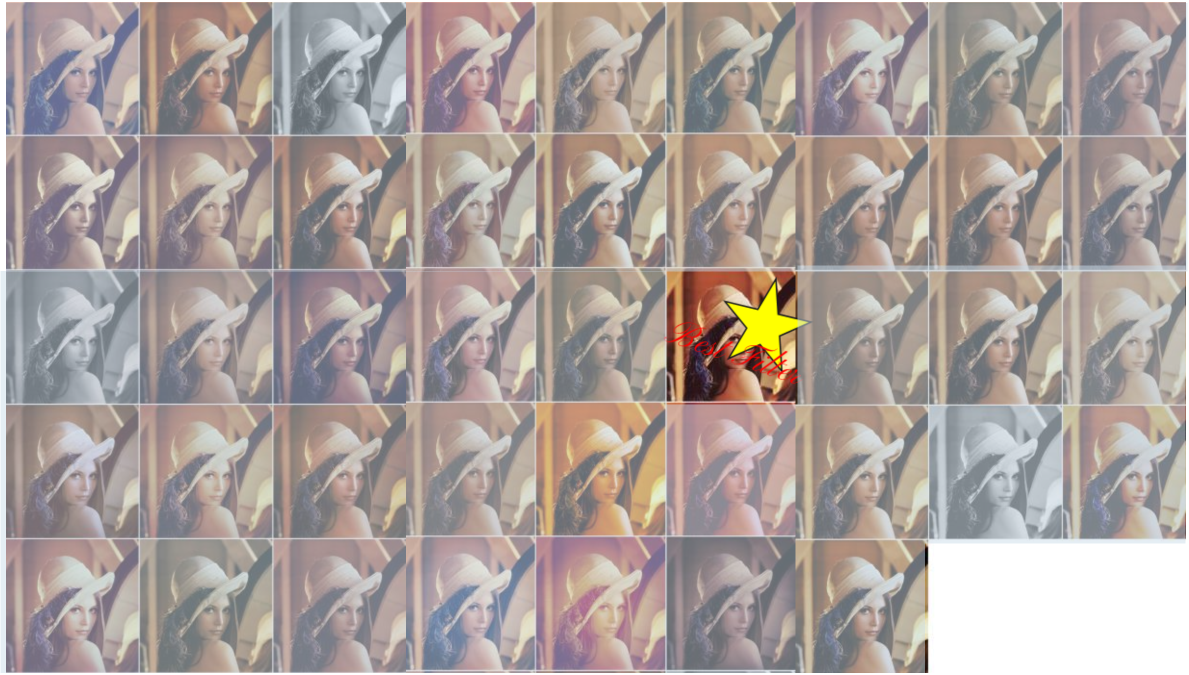


Photo Filter Recommendation



Filter recommendation: Challenges

- A subjective task
- Collecting a large corpus of labeled data for training a recommendation approach is highly expensive.

Contributions

- Demonstrate how filtered images on Instagram can be used for training a filter recommendation model
- Use high level image representations for predicting proper filters
- Achieve 51.87% top-1 accuracy on FACD

Outline

- Introduction
- **Approach**
  - Instagram Filtered Image Dataset
  - Filter Recommendation Network
- Experiments
- Conclusion

Instagram Filtered Image Dataset

- 19 filters
- 68,400 photos (3,600 photos per filter)
- No photo duplication
- No noisy filter categorization (manual inspection)

Filter Aesthetic Comparison Dataset (FACD)

- 1,280 reference images
  - 1,120 images for training
  - 160 images for evaluation
- 28,160 filtered images created from 22 filters
- 42,240 filtered image pairs with aesthetic comparison labels

W. -T. Sun, T. -H. Chao, Y. -H. Kuo, Winston H. Hsu. Photo Filter Recommendation by Category-Aware Aesthetic Learning. *IEEE Trans. on Multimedia*, 2017.

Instagram vs. FACD

	Pros	Cons
Instagram dataset	• Low cost of collection • High diversity	• **No original photo** • One filter per photo (decided by user)
FACD	• Original photo available • Multiple filters per photo (decided by AMT)	• High cost of ground truth construction • Low quantity

Selection of photo filters depends on photo content

Food

Soft color

1977, Aden, Sutro

Selfie

Bright color

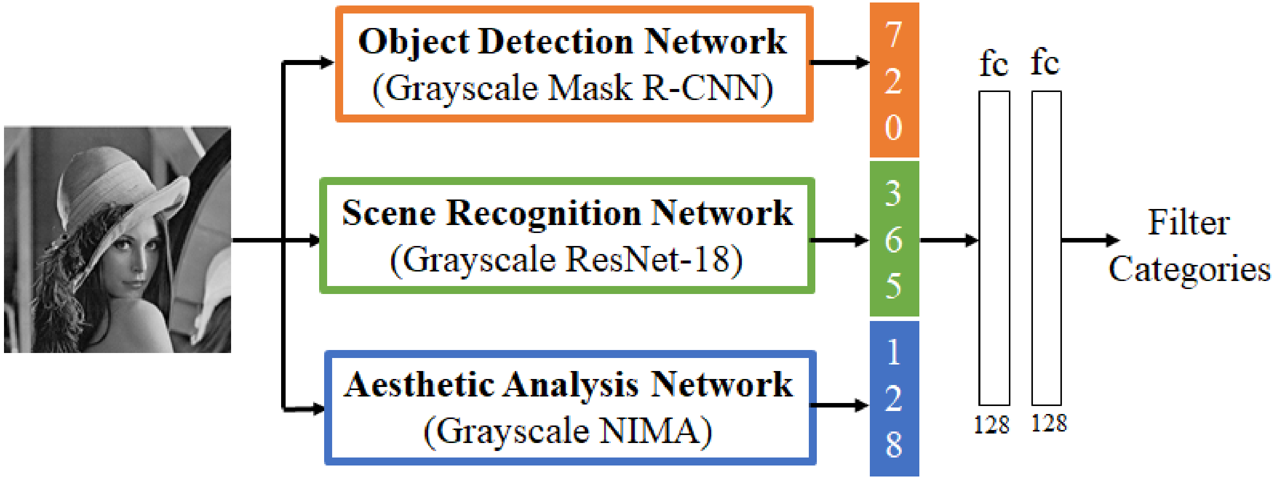
Hefe, Slumber

Natural scene

Strong contrast

Amaro, Brooklyn

Filter Recommendation Network



Why grayscale inputs?

- Training images are *filtered images*.
- Filter recommendation vs. Filter categorization
- Extract ***style-invariant*** image features

Object Detection Network

- Mask R-CNN + MS COCO (80 classes)

Model	AP^{bb}	AP^{bb}_{50}	AP^{bb}_{75}
Mask R-CNN	38.2	60.3	41.7
Grayscale Mask R-CNN	33.9	51.4	37.4

[Mask R-CNN] Kaiming He, Georgia Gkioxari, Piotr Dollár, Ross Girshick. Mask R-CNN. arXiv:1703.06870, 2017.

[MS COCO] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, Piotr Dollár. Microsoft COCO: Common Objects in Context. *ECCV*, 2014.

Scene Recognition Network

- Resnet-18 + Places365

Model	Top-1 acc.	Top-5 acc.
ResNet-18	54.74%	85.08%
Grayscale ResNet-18	51.00%	82.00%

[Resnet] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Deep Residual Learn-ing for Image Recognition. arXiv:1512.03385, 2015.
[Places365] B. Zhou, A. Lapedriza, J. Xiao, A. Torralba, and A. Oliva. Learning Deep Features for Scene Recognition using Places Database. *NIPS*, 2014.

Aesthetic Analysis Network

- NIMA + AVA

Model	Top-1 acc.
NIMA	57.84%
Grayscale NIMA	56.43%

[NIMA] Hossein Talebi, Peyman Milanfar. NIMA: Neural Image Assessment. arXiv:1709.054242017.
[AVA] Naila Murray, Luca Marchesotti, Florent Perronnin. AVA: A large-scale da-tabase for aesthetic visual analysis. *CVPR*, 2012.

From Image Representations to Filter Category

- Two fully connected layers, with 128 neurons in each layer
- Batch normalization
- Dropout

Outline

- Introduction
- Approach
  - Instagram Filtered Image Dataset
  - Filter Recommendation Network
- **Experiments**
- Conclusion

Setting

- FACD
  - Training: 960 (for fine-tuning the fc layers)
  - Validation: 160
  - Testing: 160
- Instagram filtered image dataset for training three sub-networks

Recommendation Results

Model	Top-1 Accuracy (%)	Top-3 Accuracy (%)
AlexNet [3]	33.13	70.63
RAPID net [4]	37.50	72.50
Category-aware learning (AlexNet) [2]	41.25	**80.00**
Category-aware learning (RAPID) [2]	41.88	79.50
Ours (Scene + Aesthetics)	**51.25**	**80.00**

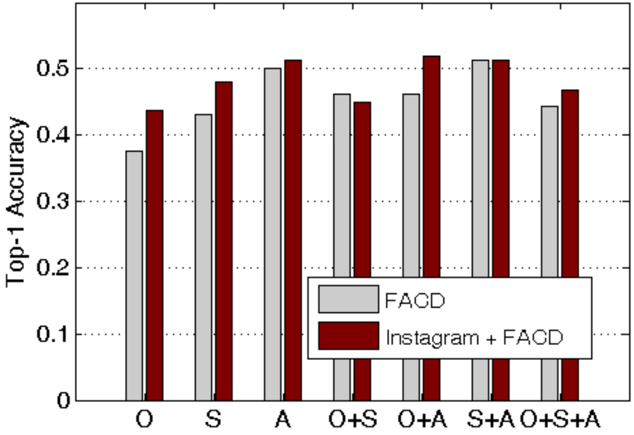
The proposed method achieved the best performances in both of top-1 and top-3 predictions. The gain was significant on the top-1 accuracy (over 9%).

Effects of Features

Model	Top-1 Accuracy (%)	Top-3 Accuracy (%)
Object	43.75	76.25
Scene	48.12	79.37
Aesthetics	51.25	75.62
Object + Scene	45.00	75.62
Object + Aesthetics	**51.87**	77.50
Scene + Aesthetics	51.25	**80.00**
Object + Scene + Aesthetics	46.87	75.00

Our approach using a single type of feature outperformed previous approaches; scenes and the aesthetics outperformed objects.

Does Instagram filtered images help?



The use of Instagram filtered images alleviated the filter recommendation task.

Outline

- Introduction
- Approach
  - Instagram Filtered Image Dataset
  - Filter Recommendation Network
- Experiments
- **Conclusion**

Conclusion

- We present a new filter recommendation method using voluminous filtered images on Instagram.
- Experimental results using the FACD benchmark dataset show the effectiveness of
  - High-level image representations
  - Using Instagram images for training the model

Photo Filter Recommendation Through Analyzing Objects, Scenes and Aesthetics

More information:
http://www2.csie.ntnu.edu.tw/~myeh